

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Wilson, Peter J
Morris, Charles P
Anson, Donald S
Occhiadoro, Teresa
Bielicki, Julie
Clements, Peter R
Hopwood, John J

(ii) TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
IDURONATE 2-SULFATASE

(iii) NUMBER OF SEQUENCES: 15

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Scully, Scott, Murphy & Presser
(B) STREET: 400 Garden City Plaza
(C) CITY: Garden City
(D) STATE: New York
(E) COUNTRY: USA
(F) ZIP: 11530

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 991,973
(B) FILING DATE: 17-DEC-1992
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: DiGiglio, Frank S
(B) REGISTRATION NUMBER: 31,346
(C) REFERENCE/DOCKET NUMBER: 8416Z

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 516-742-4343
(B) TELEFAX: 516-742-4366
(C) TELEX: 230 901 SANS UR

44

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2297 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 125..1774

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CGGCTGTGTT	GCGCAGTCTT	CATGGGTTCC	CGACGAGGAG	GTCTCTGTGG	CTGCGGCGGC	60
TGCTAACTGC	GCCACCTGCT	GCAGCCTGTC	CCCGCCGCTC	TGAAGCGGCC	GCGTCGAAGC	120
CGAA ATG CCG CCA CCC CGG ACC GGC CGA GGC CTT CTC TGG CTG GGT CTG	Met Pro Pro Pro Arg Thr Gly Arg Gly Leu Leu Trp Leu Gly Leu					169
1 5 10 15						
GTG CTG AGC TCC GTC TGC GTC GCC CTC GGA TCC GAA ACG CAG GCC AAC	Val Leu Ser Ser Val Cys Val Ala Leu Gly Ser Glu Thr Gln Ala Asn					217
20 25 30						
TCG ACC ACA GAT GCT CTG AAC GTT CTT CTC ATC ATC GTG GAT GAC CTG	Ser Thr Asp Ala Leu Asn Val Leu Leu Ile Ile Val Asp Asp Leu					265
35 40 45						
CGC CCC TCC CTG GGC TGT TAT GGG GAT AAG CTG GTG AGG TCC CCA AAT	Arg Pro Ser Leu Gly Cys Tyr Gly Asp Lys Leu Val Arg Ser Pro Asn					313
50 55 60						
ATT GAC CAA CTG GCA TCC CAC AGC CTC CTC TTC CAG AAT GCC TTT GCG	Ile Asp Gln Leu Ala Ser His Ser Leu Leu Phe Gln Asn Ala Phe Ala					361
65 70 75						
CAG CAA GCA GTG TGC GCC CCG AGC CGC GTT TCT TTC CTC ACT GGC AGG	Gln Gln Ala Val Cys Ala Pro Ser Arg Val Ser Phe Leu Thr Gly Arg					409
80 85 90 95						
AGA CCT GAC ACC ACC CGC CTG TAC GAC TTC AAC TCC TAC TGG AGG GTG	Arg Pro Asp Thr Thr Arg Leu Tyr Asp Phe Asn Ser Tyr Trp Arg Val					457
100 105 110						
CAC GCT GGA AAC TTC TCC ACC ATC CCC CAG TAC TTC AAG GAG AAT GGC	His Ala Gly Asn Phe Ser Thr Ile Pro Gln Tyr Phe Lys Glu Asn Gly					505
115 120 125						

45

TAT	GTG	ACC	ATG	TCG	GTG	GGA	AAA	GTC	TTT	CAC	CCT	GGG	ATA	TCT	TCT	553
Tyr	Val	Thr	Met	Ser	Val	Gly	Lys	Val	Phe	His	Pro	Gly	Ile	Ser	Ser	
130					135					140						
AAC	CAT	ACC	GAT	GAT	TCT	CCG	TAT	AGC	TGG	TCT	TTT	CCA	CCT	TAT	CAT	601
Asn	His	Thr	Asp	Asp	Ser	Pro	Tyr	Ser	Trp	Ser	Phe	Pro	Pro	Tyr	His	
145					150					155						
CCT	TCC	TCT	GAG	AAG	TAT	GAA	AAA	ACT	AAG	ACA	TGT	CGA	GGG	CCA	GAT	649
Pro	Ser	Ser	Glu	Lys	Tyr	Glu	Asn	Thr	Lys	Thr	Cys	Arg	Gly	Pro	Asp	
160					165					170				175		
GGA	GAA	CTC	CAT	GCC	AAC	CTG	CTT	TGC	CCT	GTG	GAT	GTG	CTG	GAT	GTT	697
Gly	Glu	Leu	His	Ala	Asn	Leu	Leu	Cys	Pro	Val	Asp	Val	Leu	Asp	Val	
180									185					190		
CCC	GAG	GGC	ACC	TTG	CCT	GAC	AAA	CAG	AGC	ACT	GAG	CAA	GCC	ATA	CAG	745
Pro	Glu	Gly	Thr	Leu	Pro	Asp	Lys	Gln	Ser	Thr	Glu	Gln	Ala	Ile	Gln	
195								200						205		
TTG	TTG	GAA	AAG	ATG	AAA	ACG	TCA	GCC	AGT	CCT	TTC	TTC	CTG	GCC	GTT	793
Leu	Leu	Glu	Lys	Met	Lys	Thr	Ser	Ala	Ser	Pro	Phe	Phe	Leu	Ala	Val	
210							215						220			
GCG	TAT	CAT	AAG	CCA	CAC	ATC	CCC	TTC	AGA	TAC	CCC	AAG	GAA	TTT	CAG	841
Gly	Tyr	His	Lys	Pro	His	Ile	Pro	Phe	Arg	Tyr	Pro	Lys	Glu	Phe	Gln	
225							230					235				
AAG	TTG	TAT	CCC	TTG	GAG	AAC	ATC	ACC	CTG	GCC	CCC	GAT	CCC	GAG	GTC	889
Lys	Leu	Tyr	Pro	Leu	Glu	Asn	Ile	Thr	Leu	Ala	Pro	Asp	Pro	Glu	Val	
240							245				250			255		
CCT	GAT	GGC	CTA	CCC	CCT	GTG	GCC	TAC	AAC	CCC	TGG	ATG	GAC	ATC	AGG	937
Pro	Asp	Gly	Leu	Pro	Pro	Val	Ala	Tyr	Asn	Pro	Trp	Met	Asp	Ile	Arg	
260								265					270			
CAA	CGG	GAA	GAC	GTC	CAA	GCC	TTA	AAC	ATC	AGT	GTG	CCG	TAT	GGT	CCA	985
Gln	Arg	Glu	Asp	Val	Gln	Ala	Leu	Asn	Ile	Ser	Val	Pro	Tyr	Gly	Pro	
275								280					285			
ATT	CCT	GTG	GAC	TTT	CAG	CGG	AAA	ATC	CGC	CAG	AGC	TAC	TTT	GCC	TCT	1033
Ile	Pro	Val	Asp	Phe	Gln	Arg	Lys	Ile	Arg	Gln	Ser	Tyr	Phe	Ala	Ser	
290								295					300			
GTG	TCA	TAT	TTG	GAT	ACA	CAG	GTC	GGC	CGC	CTC	TTG	AGT	GCT	TTG	GAC	1081
Val	Ser	Tyr	Leu	Asp	Thr	Gln	Val	Gly	Arg	Leu	Leu	Ser	Ala	Leu	Asp	
305								310				315				
GAT	CTT	CAG	CTG	GCC	AAC	AGC	ACC	ATC	ATT	GCA	TTT	ACC	TCG	GAT	CAT	1129
Asp	Leu	Gln	Leu	Ala	Asn	Ser	Thr	Ile	Ile	Ala	Phe	Thr	Ser	Asp	His	
320								325				330			335	

GGG TGG GCT CTA GGT GAA CAT GGA GAA TGG GCC AAA TAC AGC AAT TTT		1177
Gly Trp Ala Leu Gly Glu His Gly Glu Trp Ala Lys Tyr Ser Asn Phe		
340	345	350
GAT GTT GCT ACC CAT GTT CCC CTG ATA TTC TAT GTT CCT GGA AGG ACG		1225
Asp Val Ala Thr His Val Pro Leu Ile Phe Tyr Val Pro Gly Arg Thr		
355	360	365
GCT TCA CTT CCG GAG GCA GGC GAG AAG CTT TTC CCT TAC CTC GAC CCT		1273
Ala Ser Leu Pro Glu Ala Gly Glu Lys Leu Phe Pro Tyr Leu Asp Pro		
370	375	380
TTT GAT TCC GCC TCA CAG TTG ATG GAG CCA GGC AGG CAA TCC ATG GAC		1321
Phe Asp Ser Ala Ser Gln Leu Met Glu Pro Gly Arg Gln Ser Met Asp		
385	390	395
CTT GTG GAA CTT GTG TCT CTT TTT CCC ACG CTG GCT GGA CTT GCA GGA		1369
Leu Val Glu Leu Val Ser Leu Phe Pro Thr Leu Ala Gly Leu Ala Gly		
400	405	410
415		
CTG CAG GTT CCA CCT CGC TGC CCC GTT CCT TCA TTT CAC GTT GAG CTG		1417
Leu Gln Val Pro Pro Arg Cys Pro Val Pro Ser Phe His Val Glu Leu		
420	425	430
TGC AGA GAA GGC AAG AAC CTT CTG AAG CAT TTT CGA TTC CGT GAC TTG		1465
Cys Arg Glu Gly Lys Asn Leu Leu Lys His Phe Arg Phe Arg Asp Leu		
435	440	445
GAA GAG GAT CCG TAC CTC CCT GGT AAT CCC CGT GAA CTG ATT GCC TAT		1513
Glu Glu Asp Pro Tyr Leu Pro Gly Asn Pro Arg Glu Leu Ile Ala Tyr		
450	455	460
AGC CAG TAT CCC CGG CCT TCA GAC ATC CCT CAG TGG AAT TCT GAC AAG		1561
Ser Gln Tyr Pro Arg Pro Ser Asp Ile Pro Gln Trp Asn Ser Asp Lys		
465	470	475
CCG AGT TTA AAA GAT ATA AAG ATC ATG GGC TAT TCC ATA CGC ACC ATA		1609
Pro Ser Leu Lys Asp Ile Lys Ile Met Gly Tyr Ser Ile Arg Thr Ile		
480	485	490
495		
GAC TAT AGG TAT ACT GTG TGG GTT GGC TTC AAT CCT GAT GAA TTT CTA		1657
Asp Tyr Arg Tyr Thr Val Trp Val Gly Phe Asn Pro Asp Glu Phe Leu		
500	505	510
GCT AAC TTT TCT GAC ATC CAT GCA GGG GAA CTG TAT TTT GTG GAT TCT		1705
Ala Asn Phe Ser Asp Ile His Ala Gly Glu Leu Tyr Phe Val Asp Ser		
515	520	525
GAC CCA TTG CAG GAT CAC AAT ATG TAT AAT GAT TCC CAA GGT GGA GAT		1753
Asp Pro Leu Gln Asp His Asn Met Tyr Asn Asp Ser Gln Gly Gly Asp		
530	535	540

47

CTT TTC CAG TTG TTG ATG CCT TGAGTTTGC CAACCATGGA TGGCAAATGT	1804
Leu Phe Gln Leu Leu Met Pro	
545	550
GATGTGCTCC CTTCCAGCTG GTGAGAGGAG GAGTTAGAGC TGGTCGTTT GTGATTACCC	1864
ATAATATTGG AAGCAGCCTG AGGGCTAGTT AATCCAAACA TGCATCAACA ATTTGGCCTG	1924
AGAATATGTA ACAGCCAAAC CTTTCGTTT AGTCTTTATT AAAATTTATA ATTGGTAATT	1984
GGACCAGTTT TTTTTTAAT TTCCCTCTT TTAAAACAGT TACGGCTTAT TTACTGAATA	2044
AATACAAAGC AAACAAACTC AAGTTATGTC ATACCTTG ATACGAAGAC CATACTAAT	2104
AAACAAACAT AACATTATAC ACAAAAGAATA CTTTCATTAT TTGTGGAATT TAGTGCATTT	2164
CAAAAAGTAA TCATATATCA AACTAGGCAC CACACTAAGT TCCTGATTAT TTTGTTATA	2224
ATTTAATAAT ATATCTTATG AGCCCTATAT ATTCAAAATA TTATGTTAAC ATGTAATCCA	2284
TGTTTCTTT TCC	2297

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 550 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Pro Pro Pro Arg Thr Gly Arg Gly Leu Leu Trp Leu Gly Leu Val			
1	5	10	15
Leu Ser Ser Val Cys Val Ala Leu Gly Ser Glu Thr Gln Ala Asn Ser			
20	25	30	
Thr Thr Asp Ala Leu Asn Val Leu Leu Ile Ile Val Asp Asp Leu Arg			
35	40	45	
Pro Ser Leu Gly Cys Tyr Gly Asp Lys Leu Val Arg Ser Pro Asn Ile			
50	55	60	
Asp Gln Leu Ala Ser His Ser Leu Leu Phe Gln Asn Ala Phe Ala Gln			
65	70	75	80
Gln Ala Val Cys Ala Pro Ser Arg Val Ser Phe Leu Thr Gly Arg Arg			
85	90	95	
Pro Asp Thr Thr Arg Leu Tyr Asp Phe Asn Ser Tyr Trp Arg Val His			
100	105	110	

Ala Gly Asn Phe Ser Thr Ile Pro Gln Tyr Phe Lys Glu Asn Gly Tyr
 115 120 125

Val Thr Met Ser Val Gly Lys Val Phe His Pro Gly Ile Ser Ser Asn
 130 135 140

His Thr Asp Asp Ser Pro Tyr Ser Trp Ser Phe Pro Pro Tyr His Pro
 145 150 155 160

Ser Ser Glu Lys Tyr Glu Asn Thr Lys Thr Cys Arg Gly Pro Asp Gly
 165 170 175

Glu Leu His Ala Asn Leu Leu Cys Pro Val Asp Val Leu Asp Val Pro
 180 185 190

Glu Gly Thr Leu Pro Asp Lys Gln Ser Thr Glu Gln Ala Ile Gln Leu
 195 200 205

Leu Glu Lys Met Lys Thr Ser Ala Ser Pro Phe Phe Leu Ala Val Gly
 210 215 220

Tyr His Lys Pro His Ile Pro Phe Arg Tyr Pro Lys Glu Phe Gln Lys
 225 230 235 240

Leu Tyr Pro Leu Glu Asn Ile Thr Leu Ala Pro Asp Pro Glu Val Pro
 245 250 255

Asp Gly Leu Pro Pro Val Ala Tyr Asn Pro Trp Met Asp Ile Arg Gln
 260 265 270

Arg Glu Asp Val Gln Ala Leu Asn Ile Ser Val Pro Tyr Gly Pro Ile
 275 280 285

Pro Val Asp Phe Gln Arg Lys Ile Arg Gln Ser Tyr Phe Ala Ser Val
 290 295 300

Ser Tyr Leu Asp Thr Gln Val Gly Arg Leu Leu Ser Ala Leu Asp Asp
 305 310 315 320

Leu Gln Leu Ala Asn Ser Thr Ile Ile Ala Phe Thr Ser Asp His Gly
 325 330 335

Trp Ala Leu Gly Glu His Gly Glu Trp Ala Lys Tyr Ser Asn Phe Asp
 340 345 350

Val Ala Thr His Val Pro Leu Ile Phe Tyr Val Pro Gly Arg Thr Ala
 355 360 365

Ser Leu Pro Glu Ala Gly Glu Lys Leu Phe Pro Tyr Leu Asp Pro Phe
 370 375 380

Asp Ser Ala Ser Gln Leu Met Glu Pro Gly Arg Gln Ser Met Asp Leu
 385 390 395 400

Val Glu Leu Val Ser Leu Phe Pro Thr Leu Ala Gly Leu Ala Gly Leu
 405 410 415

Gln Val Pro Pro Arg Cys Pro Val Pro Ser Phe His Val Glu Leu Cys
 420 425 430

Arg Glu Gly Lys Asn Leu Leu Lys His Phe Arg Phe Arg Asp Leu Glu
 435 440 445

Glu Asp Pro Tyr Leu Pro Gly Asn Pro Arg Glu Leu Ile Ala Tyr Ser
 450 455 460

Gln Tyr Pro Arg Pro Ser Asp Ile Pro Gln Trp Asn Ser Asp Lys Pro
 465 470 475 480

Ser Leu Lys Asp Ile Lys Ile Met Gly Tyr Ser Ile Arg Thr Ile Asp
 485 490 495

Tyr Arg Tyr Thr Val Trp Val Gly Phe Asn Pro Asp Glu Phe Leu Ala
 500 505 510

Asn Phe Ser Asp Ile His Ala Gly Glu Leu Tyr Phe Val Asp Ser Asp
 515 520 525

Pro Leu Gln Asp His Asn Met Tyr Asn Asp Ser Gln Gly Gly Asp Leu
 530 535 540

Phe Gln Leu Leu Met Pro
 545 550

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Pro Arg Glu Leu Ile Ala Tyr Ser Asn Tyr Pro Arg Asn Asn Ile Pro
 1 5 10 15

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 49 base pairs
- (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ACTAGTAGCA CCTGCTGGAC GCCGGGAGGG ACCCGCTGAT GCTGCTGCA

49

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Thr	Ser	Ala	Leu	Asn	Val	Leu	Leu	Ile	Ile	Val	Asp	Asp	Leu	Arg	Pro
1					5									15	
Ser	Leu	Gly	Asp	Tyr	Asp	Asp	Val	Leu							
					20										25

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4428 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 332..434

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 536..537

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 693..829

51

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 962..963

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1044..1221

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 1350..1351

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1480..1569

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 1716..1717

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1841..2041

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 2206..2207

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2294..2464

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 2585..2586

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2684..2810

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 2904..2905

(ix) FEATURE:

(A) NAME/KEY: CDS
 (B) LOCATION: 3033..3206

(ix) FEATURE:

(A) NAME/KEY: intron
 (B) LOCATION: 3308..3309

(ix) FEATURE:

(A) NAME/KEY: CDS
 (B) LOCATION: 3435..3908

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GCGATCTAGA CCTAGTTAGC	CAAGTCTCTA ACGTGACATA	GGGAAAGCTT GCAATGGCAA	60		
CTGGCCGCC	GTCTGCGCCT GTCTCTCGCC	ACGCCTATTG CTGCAGGATG	ACGCGCACCT	120	
CTATGAACCC	GCCGTGAGGT GTGAGTGTGA	CGCAGGGAAG AGTCGCACGG	ACGCACTCGC	180	
G	GCTGCGGCCA GCTGCGGCC	CGGGCGGCCG	CTGTGTTGCG CAGTCTTCAT	GGGTTCCCGA	240
CGAGGAGGTC	TCTGTGGCTG CGGC GGCTGC	TAACTGCGCC ACCTGCTGCA	GCCTGTC	300	
CCC GCTCTGA	AGCGGCCGCG TCGAAGCCGA	AATGCCGCCA	CCCCGGACCG	GCCGAGGCCT	360
TCTCTGGCTG	GGTCTGGTTG TGAGCTCCGT	CTGC GTCGCC	CTCGGATCCG	AAACGCAGGC	420
CAACTCGACC	ACAGGTGCCG CCCACGCCCT	CCCTGCCATC	TCTTCTCCCT	TCCTCCCTCC	480
CTTCCTTCCT	CCTTCCTTCT TTCCCTCCTT	CTTTGTTTAT	ATCCATTCTT	TTTACCCCCC	540
ACTCCCACCC	TTGCTGAGGC ACAGCGCCCT	CCCTGGCTAG	GCTGTTAGGT	GCAGGGTCCA	600
GCCTTGGGCC	TCTTAGTAAC CTAGCACCTA	CCATGAGGGA	GGGTTCA GTG	TCAGTGCAGG	660
TTACCTCACC	AAAGCCCCTC CCTCCTGTGT	AGATGCTCTG	AACGTTCTTC	TCATCATCGT	720
GGATGACCTG	CGCCCCTCCC TGGGCTGTTA	TGGGGATAAG	CTGGTGAGGT	CCCCAAATAT	780
TGACCAACTG	GCATCCCACA GCCTCCTCTT	CCAGAATGCC	TTTGGCAGG	TATGTCTGGG	840
AACCTCTAGC	TGTGGGTGTG TGCTGCTTCG	TGCACTGAGG	GTTGGGGCG	GGGAGCTTCA	900
GCTATTGTCA	GATGGCACAG ATTGTGCGGG	ACATCTGTT	AGAGGGAAGC	ATAGTCTGGA	960
AAAGGGCGGT	TGCTTGGTTA CCTAAGAGAT	GGCAGACATG	TTTGCTGTG	GCGATGCTTA	1020
CCTCTGCTTC	TGCTCCCTAA CAGCAAGCAG	TGTGCGCCCC	GAGCCGCGTT	TCTTTCCCTCA	1080
CTGGCAGGAG	ACCTGACACC ACCCGCCTGT	ACGACTTCAA	CTCCTACTGG	AGGGTGCACG	1140

CTGGAAACTT	CTCCACCATC	CCCCAGTACT	TCAAGGAGAA	TGGCTATGTG	ACCATGTCGG	1200
TGGGAAAAGT	CTTCACCCCT	GGTACTGCTC	CATGTCCAGA	GTCTGGTTTC	TCTTGGTTTG	1260
TGGTGTCTGA	NTCCAGCATT	CCCATCCTGG	GGATGGGCTG	TCTTGCAGA	GCCCTCTTCT	1320
GGCTGGCGA	GTCCCTCGCT	AGTCAGTGCT	TTTGTAGATG	AGGAAACTGA	GCCCCAAAGA	1380
AGGGAGGNTC	CACTTGCCCA	TTTGTAAAC	GAGTTTAAT	TATGGGGAGT	GGGGTGTGAA	1440
AAGACTCATC	ATGTTTAAC	AACTTTTT	TTTTCCAAG	GGATATCTTC	TAACCATACC	1500
GATGATTCTC	CGTATAGCTG	GTCTTTCCA	CCTTATCATC	CTTCCTCTGA	GAAGTATGAA	1560
AACACTAAGG	TAAGGCTGTG	AAAGGGACAT	TTCTGAAGAG	GAACCACTT	TTCCCTTGTC	1620
ACATAAACTA	CTGGGTATAC	TGCATGTNCT	GTGAAGCTGG	TTATATACCA	CGAAGTTGTG	1680
GGTTTCATT	GTGATAATGT	TTTGACAGAA	GTAAGTTGTT	CAGTCTGAGT	GACTAACACG	1740
TGAAGGGCTG	ATTATGTGAA	CATTAAATCT	GTGTGTGTAG	CCTTCATGGC	TTCATNTCTT	1800
GCACCTAAAA	AGCTGATGTT	ATATTATTT	TTTTGAAAG	ACATGTCGAG	GGCCAGATGG	1860
AGAACTCCAT	GCCAACCTGC	TTGCCCTGT	GGATGTGCTG	GATGTTCCCG	AGGGCACCTT	1920
GCCTGACAAA	CAGAGCACTG	AGCAAGCCAT	ACAGTTGTTG	GAAAAGATGA	AAACGTCAGC	1980
CAGTCCTTTC	TTCCTGGCCG	TTGGGTATCA	TAAGCCACAC	ATCCCCTTCA	GATAACCCAA	2040
GGTGAAGAGC	TGGTTGAGGG	CTGATCCAGC	ACAGCTGTGA	CAGCTGTGTT	GTTTGTGAG	2100
GGAGGGATT	GCACAGGGAA	GGTGGCTACA	TCCTGCCATC	GCCAGGCACC	ATGGTTGCCT	2160
GATGGGCACT	AGTGTCTCA	GTGGAGTAA	GATGGGATT	AGAGGTAAAA	GGCAGTATAG	2220
ACAGTGATAG	AGCCACAAGC	TTGTGCTTT	GCTAAAAGAG	TGACAACCTT	GTGGCTTGTT	2280
GTTTTTCCCC	AAGGAATTTC	AGAAGTTGTA	TCCCTGGAG	AACATCACCC	TGGCCCCCGA	2340
TCCCGAGGTC	CCTGATGGCC	TACCCCTGT	GGCCTACAAAC	CCCTGGATGG	ACATCAGGCA	2400
ACGGGAAGAC	GTCCAAGCCT	TAAACATCAG	TGTGCCGTAT	GGTCCAATT	CTGTGGACTT	2460
TCAGGTATCA	AGGACATAGT	TTGGGGATGT	ATTGGACACT	GATGACATAG	TGTCGTAGGT	2520
GAAACCAC	TTCTCAGTAG	ACACAACTCC	ACCTATAATG	TCTTATTAAG	AGCTTTCTTT	2580
GTGTGTAGGG	ATTGGGAGAG	ATGCACACGG	CAAGCATTAT	CTCTGTATGC	CTTGGCAATT	2640
TAAATTGCAG	TCACTCTCAT	TTTATTTT	TTCAATTG	CAGCGGAAAA	TCCGCCAGAG	2700
CTACTTGCC	TCTGTGTCA	ATTGGATAC	ACAGGTCGGC	CGCCTCTTGA	GTGCTTTGGA	2760

CGATCTTCAG	CTGGCCAACA	GCACCATCAT	TGCATTTACC	TCGGATCATG	GTAAGCATTT	2820
TGAAATTCCC	TGGTGAGTCA	AAACATCTGA	ACTTCTGT	GAAACATGCT	TTGCAAAATT	2880
GCCATTGACA	TAAACATGGG	TGTGTTCTT	CTAGGTGATG	AGTTTCTACT	TCCTCTGGTT	2940
TTTACAACAG	GAAATGAAAT	GGTATCTAAA	ATAAACAAAGC	TGTGGTATGA	TGATTATTCA	3000
TTTTCTGTCA	TTCTGTGCTT	TTTATGAAC	AGGGTGGGCT	CTAGGTGAAC	ATGGAGAATG	3060
GGCCAAATAC	AGCAATTTG	ATGTTGCTAC	CCATGTTCCC	CTGATATTCT	ATGTTCTGG	3120
AAGGACGGCT	TCACTTCCGG	AGGCAGGCAG	GAAGCTTTTC	CCTTACCTCG	ACCCTTTGA	3180
TTCCGCCTCA	CAGTTGATGG	AGCCAGGTAT	AAAATATGCT	GAAATGATAT	TGCTTGACAG	3240
TAAGATCACC	TTTAGTTTAT	ATGTGAACCA	CTTATTGAA	TCATAGGCTT	TGGGGTTACA	3300
CAGACCCCAA	AGATAAATGG	TGTAAATTAA	AAAAAGAAAA	CATATGGAGC	CCAGACAGGG	3360
TCCTTTACTG	CTCCTGCCTG	GCCATGGCAG	GCTTTTATAA	TGTAACCCAT	TCTGCTCTGT	3420
CGCTTCCTGT	TTCAGGCAGG	CAATCCATGG	ACCTTGTGGA	ACTTGTGTCT	CTTTTTCCCA	3480
CGCTGGCTGG	ACTTGCAGGA	CTGCAGGTTTC	CACCTCGCTG	CCCCGTTCCCT	TCATTTCACG	3540
TGAGCTGTG	CAGAGAACGGC	AAGAACCTTC	TGAAGCATTT	TCGATTCCGT	GACTTGGAAAG	3600
AGGATCCGTA	CCTCCCTGGT	AATCCCCGTG	AACTGATTGC	CTATAGCCAG	TATCCCCGGC	3660
CTTCAGACAT	CCCTCAGTGG	AATTCTGACA	AGCCGAGTTT	AAAAGATATA	AAGATCATGG	3720
GCTATTCCAT	ACGCACCATA	GACTATAGGT	ATACTGTGTG	GGTTGGCTTC	AATCCTGATG	3780
AATTCTAGC	TAACTTTCT	GACATCCATG	CAGGGAACT	GTATTTGTG	GATTCTGACC	3840
CATTGCAGGA	TCACAATATG	TATAATGATT	CCCAAGGTGG	AGATCTTTTC	CAGTTGTTGA	3900
TGCCTTGAGT	TTTGCCAACC	ATGGATGGCA	AATGTGATGT	GCTCCCTTCC	AGCTGGTGAG	3960
AGGAGGAGTT	AGAGCTGGTC	GTGGATGGAT	TACCCATAAT	ATTGGAAGCA	GCCTGAGGGC	4020
TAGTTAATCC	AAACATGCAT	CAACAATTG	GCCTGAGAAT	ATGTAACAGC	CAAACCTTT	4080
CGTTTAGTCT	TTATTAAAAT	TTATAATTGG	TAATTGGACC	AGTTTTTTT	TTAATTCCC	4140
TCTTTTAAA	ACAGTTACGG	CTTATTTACT	GAATAAATAC	AAAGCAAACA	AACTCAAGTT	4200
ATGTCATACC	TTTGGATACG	AAGACCATAAC	ATAATAACCA	AACATAACAT	TATACACAAA	4260
GAATACTTTC	ATTATTTGTG	GAATTTAGTG	CATTTCAAAA	AGTAATCATA	TATCAAACTA	4320

GGCACACAC TAAGTTCTTG ATTATTTGT TTATAATTAA ATAATATATC TTATGAGCCC 4380
 TATATATTCA AAATATTATG TTAACATGTA ATCCATGTTT CTTTTCC 4428

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Pro Arg Glu Leu Ile Ala Tyr Ser Xaa Tyr Pro Arg Xaa Xaa Ile Pro
 1 5 10 15

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Cys Thr Pro Ser Arg
 1 5

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Gly Lys Trp His Leu Gly
 1 5

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 81 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGCCTCTAGA CCAGCTACAG TCGGAAACCA TCAGCAAGCA GGTCAATTGTT CCAACATGCC

60

GCCACCCCGG ACCGGCCGAG G

81

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 510 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Thr	Arg	Arg	Pro	Asn	Val	Val	Leu	Leu	Leu	Thr	Asp	Asp	Gln	Asp	Glu
1					5					10			15		

Val	Leu	Gly	Gly	Met	Thr	Pro	Leu	Lys	Lys	Thr	Lys	Ala	Leu	Ile	Gly
				20				25				30			

Glu	Met	Gly	Met	Thr	Phe	Ser	Ser	Ala	Tyr	Val	Pro	Ser	Ala	Leu	Cys
	35					40						45			

Cys	Pro	Ser	Arg	Ala	Ser	Ile	Leu	Thr	Gly	Lys	Tyr	Pro	His	Asn	His
	50					55				60					

His	Val	Val	Asn	Asn	Thr	Leu	Glu	Gly	Asn	Cys	Ser	Ser	Lys	Ser	Trp
	65				70				75				80		

Gln	Lys	Ile	Gln	Glu	Pro	Asn	Thr	Phe	Pro	Ala	Ile	Leu	Arg	Ser	Met
	85						90						95		

Gln	Gly	Tyr	Gln	Thr	Phe	Thr	Phe	Phe	Ala	Gly	Lys	Tyr	Leu	Asn	Glu
	100					105						110			

Tyr	Gly	Ala	Pro	Asp	Ala	Gly	Gly	Leu	Glu	His	Val	Pro	Leu	Gly	Trp
	115					120					125				

Ser Tyr Trp Tyr Ala Leu Glu Lys Asn Ser Lys Tyr Tyr Asn Tyr Thr
 130 135 140

Leu Ser Ile Asn Gly Lys Ala Arg Lys His Gly Glu Asn Tyr Ser Val
 145 150 155 160

Asp Tyr Leu Thr Asp Val Leu Ala Asn Val Ser Leu Asp Phe Leu Asp
 165 170 175

Tyr Lys Ser Asn Glu Glu Pro Phe Phe Met Met Ile Ala Thr Pro Ala
 180 185 190

Pro His Ser Pro Trp Thr Ala Ala Pro Gln Tyr Gln Lys Ala Phe Gln
 195 200 205

Asn Val Phe Ala Pro Arg Asn Lys Asn Phe Asn Ile His Gly Thr Asn
 210 215 220

Lys His Trp Leu Ile Arg Gln Ala Lys Thr Pro Met Thr Asn Ser Ser
 225 230 235 240

Ile Gln Phe Leu Asp Asn Ala Phe Arg Lys Arg Trp Gln Thr Leu Leu
 245 250 255

Ser Val Asp Asp Leu Val Glu Lys Leu Val Lys Arg Leu Glu Phe Thr
 260 265 270

Gly Glu Leu Asn Asn Thr Tyr Ile Phe Tyr Thr Ser Asp Asn Gly Tyr
 275 280 285

His Thr Gly Gln Phe Ser Leu Pro Ile Asp Lys Arg Gln Leu Tyr Glu
 290 295 300

Phe Asp Ile Lys Val Pro Leu Leu Val Arg Gly Pro Gly Ile Lys Pro
 305 310 315 320

Asn Gln Thr Ser Lys Met Leu Val Ala Asn Ile Asp Leu Gly Pro Ile
 325 330 335

Leu Asp Ile Ala Gly Tyr Asp Leu Asn Lys Thr Gln Met Asp Gly Met
 340 345 350

Ser Leu Leu Pro Ile Leu Arg Gly Ala Ser Asn Leu Thr Trp Arg Ser
 355 360 365

Asp Val Leu Val Glu Tyr Gln Gly Glu Gly Arg Asn Val Thr Asp Pro
 370 375 380

Thr Cys Pro Ser Leu Ser Pro Gly Val Ser Gln Cys Phe Pro Asp Cys
 385 390 395 400

Val Cys Glu Asp Ala Tyr Asn Asn Thr Tyr Ala Cys Val Arg Thr Met
 405 410 415

Ser Ala Leu Trp Asn Leu Gln Tyr Cys Glu Phe Asp Asp Gln Glu Val
 420 425 430
 Phe Val Glu Val Tyr Asn Leu Thr Ala Asp Pro Asp Gln Ile Thr Asn
 435 440 445
 Ile Ala Lys Thr Ile Asp Pro Glu Leu Leu Gly Lys Met Asn Tyr Arg
 450 455 460
 Leu Met Met Leu Gln Ser Cys Ser Gly Pro Thr Cys Arg Thr Pro Gly
 465 470 475 480
 Val Phe Asp Pro Gly Tyr Arg Phe Asp Pro Arg Leu Met Phe Ser Asn
 485 490 495
 Arg Gly Ser Val Arg Thr Arg Arg Phe Ser Lys His Leu Leu
 500 505 510

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 507 amino acids.
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Gly Ala Pro Arg Ser Leu Leu Leu Ala Leu Ala Ala Gly Leu Ala
 1 5 10 15
 Val Ala Arg Pro Pro Asn Ile Val Leu Ile Phe Ala Asp Asp Leu Gly
 20 25 30
 Tyr Gly Asp Leu Gly Cys Tyr Gly His Pro Ser Ser Thr Thr Pro Asn
 35 40 45
 Leu Asp Gln Leu Ala Ala Gly Gly Leu Arg Phe Thr Asp Phe Tyr Val
 50 55 60
 Pro Val Ser Leu Gln Thr Pro Ser Arg Ala Ala Leu Leu Thr Gln Arg
 65 70 75 80
 Leu Pro Val Arg Met Gly Met Tyr Pro Gly Val Leu Val Pro Ser Ser
 85 90 95
 Arg Gly Gly Leu Pro Leu Glu Glu Val Thr Val Ala Glu Val Leu Ala
 100 105 110
 Ala Arg Gly Tyr Leu Thr Gly Met Ala Gly Lys Trp His Leu Gly Val
 115 120 125

Gly Pro Glu Gly Ala Phe Leu Pro Pro His Gln Gly Phe His Arg Phe
 130 135 140

Leu Gly Ile Pro Tyr Ser His Asp Gln Gly Pro Cys Gln Asn Leu Thr
 145 150 155 160

Cys Phe Pro Pro Ala Thr Pro Cys Asp Gly Gly Cys Asp Gln Gly Leu
 165 170 175

Val Pro Ile Pro Leu Leu Ala Asn Leu Ser Val Glu Ala Gln Pro Pro
 180 185 190

Trp Leu Pro Gly Leu Glu Ala Arg Tyr Met Ala Phe Ala His Asp Leu
 195 200 205

Met Ala Asp Ala Gln Arg Gln Asp Arg Pro Phe Phe Leu Tyr Tyr Ala
 210 215 220

Ser His His Thr His Tyr Pro Gln Phe Ser Gly Gln Ser Phe Ala Glu
 225 230 235 240

Arg Ser Gly Arg Gly Pro Phe Gly Asp Ser Leu Met Glu Leu Asp Ala
 245 250 255

Ala Val Gly Thr Leu Met Thr Ala Ile Gly Asp Leu Gly Leu Leu Glu
 260 265 270

Glu Thr Leu Val Ile Phe Thr Ala Asp Asn Gly Pro Glu Thr Met Arg
 275 280 285

Met Ser Arg Gly Gly Cys Ser Gly Leu Leu Arg Cys Gly Lys Gly Thr
 290 295 300

Thr Tyr Glu Gly Gly Val Arg Glu Pro Ala Leu Ala Phe Trp Pro Gly
 305 310 315 320

His Ile Ala Pro Gly Val Thr His Glu Leu Ala Ser Ser Leu Asp Leu
 325 330 335

Leu Pro Thr Leu Ala Ala Leu Ala Gly Ala Pro Leu Pro Asn Val Thr
 340 345 350

Leu Asp Gly Phe Asp Leu Arg Pro Pro Ala Ala Gly His Arg Gln Glu
 355 360 365

Pro Ser Ala Val Ser Leu Leu Leu Pro Val Leu Pro Arg Arg Gly Pro
 370 375 380

Trp Gly Phe Cys Cys Ala Asp Trp Lys Val Gln Gly Ser Leu Leu His
 385 390 395 400

Pro Gly Ser Ala His Ser Asp Thr Thr Ala Asp Pro Ala Cys His Ala
 405 410 415

60

Ser Ser Ser Leu Thr Ala His Glu Pro Pro Leu Leu Tyr Asp Leu Ser
 420 425 430
 Lys Asp Pro Gly Glu Asn Tyr Asn Leu Leu Gly Gly Val Ala Gly Ala
 435 440 445
 Thr Pro Glu Val Leu Gln Ala Leu Lys Gln Leu Gln Leu Leu Lys Ala
 450 455 460
 Gln Leu Asp Ala Ala Val Thr Phe Gly Pro Ser Gln Val Ala Arg Gly
 465 470 475 480
 Glu Asp Pro Ala Leu Gln Ile Cys Cys His Pro Gly Cys Thr Pro Arg
 485 490 495
 Pro Ala Cys Cys His Cys Pro Asp Pro His Ala
 500 505

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 533 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met Gly Pro Arg Gly Ala Ala Ser Leu Pro Arg Gly Pro Gly Pro Arg
 1 5 10 15

Arg Leu Leu Leu Pro Val Val Leu Pro Leu Leu Leu Leu Leu Leu
 20 25 30

Ala Pro Pro Gly Ser Gly Ala Gly Ala Ser Arg Pro Pro His Leu Val
 35 40 45

Phe Leu Leu Ala Asp Asp Leu Gly Trp Asn Asp Val Gly Phe His Gly
 50 55 60

Ser Arg Ile Arg Thr Pro His Leu Asp Ala Leu Ala Gly Gly Val
 65 70 75 80

Leu Leu Asp Asn Tyr Tyr Thr Gln Pro Leu Cys Thr Pro Ser Arg Ser
 85 90 95

Gln Leu Leu Thr Gln Arg Tyr Gln Ile Arg Thr Gly Leu Gln His Gln
 100 105 110

Ile Ile Trp Pro Cys Gln Pro Ser Cys Val Pro Leu Asp Glu Lys Leu
 115 120 125

Cl

Leu Pro Gln Leu Leu Lys Glu Ala Gly Tyr Thr Thr His Met Val Gly
 130 135 140
 Lys Trp His Leu Gly Met Tyr Arg Lys Glu Cys Leu Pro Thr Arg Arg
 145 150 155 160
 Gly Phe Asp Thr Tyr Phe Gly Tyr Leu Leu Gly Ser Glu Asp Tyr Tyr
 165 170 175
 Ser His Glu Arg Cys Thr Leu Ile Asp Ala Leu Asn Val Thr Arg Cys
 180 185 190
 Ala Leu Asp Phe Arg Asp Gly Glu Glu Val Ala Thr Gly Tyr Lys Asn
 195 200 205
 Met Tyr Ser Thr Asn Ile Phe Thr Lys Arg Ala Ile Ala Leu Ile Thr
 210 215 220
 Asn His Pro Pro Glu Lys Pro Leu Phe Leu Tyr Leu Ala Leu Gln Ser
 225 230 235 240
 Val His Glu Pro Leu Gln Val Pro Glu Glu Tyr Leu Lys Pro Tyr Asp
 245 250 255
 Phe Ile Gln Asp Lys Asn Arg His His Tyr Ala Gly Met Val Ser Leu
 260 265 270
 Met Asp Glu Ala Val Gly Asn Val Thr Ala Ala Leu Lys Ser Ser Gly
 275 280 285
 Leu Trp Asn Asn Ile Val Phe Ile Phe Ser Thr Asp Asn Gly Gly Gln
 290 295 300
 Thr Leu Ala Gly Gly Asn Asn Trp Pro Leu Arg Gly Arg Lys Trp Ser
 305 310 315 320
 Leu Trp Glu Gly Gly Val Arg Gly Val Gly Phe Val Ala Ser Pro Leu
 325 330 335
 Leu Lys Gln Lys Gly Val Lys Asn Arg Glu Leu Ile His Ile Ser Asp
 340 345 350
 Trp Leu Pro Thr Leu Val Lys Leu Ala Arg Gly His Thr Asn Gly Thr
 355 360 365
 Lys Pro Leu Asp Gly Phe Asp Val Trp Lys Thr Ile Ser Glu Gly Ser
 370 375 380
 Pro Ser Pro Arg Ile Glu Leu Leu His Asn Ile Asp Pro Asn Phe Val
 385 390 395 400
 Asp Ser Ser Pro Cys Pro Arg Asn Ser Met Ala Pro Ala Lys Asp Asp
 405 410 415

Ser Ser Leu Pro Glu Tyr Ser Ala Phe Asn Thr Ser Val His Ala Ala
 420 425 430
 Ile Arg His Gly Asn Trp Lys Leu Leu Thr Gly Tyr Pro Gly Cys Gly
 435 440 445
 Tyr Trp Phe Pro Pro Pro Ser Gln Tyr Asn Val Ser Glu Ile Pro Ser
 450 455 460
 Ser Asp Pro Pro Thr Lys Thr Leu Trp Leu Phe Asp Ile Asp Arg Asp
 465 470 475 480
 Pro Glu Glu Arg His Asp Leu Ser Arg Glu Tyr Pro His Ile Val Thr
 485 490 495
 Lys Leu Leu Ser Arg Leu Gln Phe Tyr His Lys His Ser Val Pro Val
 500 505 510
 Tyr Phe Pro Ala Gln Asp Pro Arg Cys Asp Pro Lys Ala Thr Gly Val
 515 520 525
 Trp Gly Pro Trp Met
 530

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 583 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Pro Leu Arg Lys Met Lys Ile Pro Phe Leu Leu Leu Phe Phe Leu
 1 5 10 15
 Trp Glu Ala Glu Ser His Ala Ala Ser Arg Pro Asn Ile Ile Leu Val
 20 25 30
 Met Ala Asp Asp Leu Gly Ile Gly Asp Pro Gly Cys Tyr Gly Asn Lys
 35 40 45
 Thr Ile Arg Thr Pro Asn Ile Asp Arg Leu Ala Ser Gly Gly Val Lys
 50 55 60
 Leu Thr Gln His Leu Ala Ala Ser Pro Leu Cys Ile Pro Ser Arg Ala
 65 70 75 80
 Ala Phe Met Thr Gly Arg Tyr Pro Val Arg Ser Gly Met Ala Ser Trp
 85 90 95

Ser Arg Thr Gly Val Phe Leu Phe Thr Ala Ser Ser Gly Gly Leu Pro
 100 105 110
 Thr Asp Glu Ile Thr Phe Ala Lys Leu Leu Lys Asp Gln Gly Tyr Ser
 115 120 125
 Thr Ala Leu Ile Gly Lys Trp His Leu Gly Met Ser Cys His Ser Lys
 130 135 140
 Thr Asp Phe Cys His His Pro Leu His His Gly Phe Asn Tyr Phe Tyr
 145 150 155 160
 Gly Ile Ser Leu Thr Asn Leu Arg Asp Cys Lys Pro Gly Glu Gly Ser
 165 170 175
 Val Phe Thr Thr Gly Phe Lys Arg Leu Val Phe Leu Pro Leu Gln Ile
 180 185 190
 Val Gly Val Thr Leu Leu Thr Leu Ala Ala Leu Asn Cys Leu Gly Leu
 195 200 205
 Leu His Val Pro Leu Gly Val Phe Phe Ser Leu Leu Phe Leu Ala Ala
 210 215 220
 Leu Ile Leu Thr Leu Phe Leu Gly Phe Leu His Tyr Phe Arg Pro Leu
 225 230 235 240
 Asn Cys Phe Met Met Arg Asn Tyr Glu Ile Ile Gln Gln Pro Met Ser
 245 250 255
 Tyr Asp Asn Leu Thr Gln Arg Leu Thr Val Glu Ala Ala Gln Phe Ile
 260 265 270
 Gln Arg Asn Thr Glu Thr Pro Phe Leu Leu Val Leu Ser Tyr Leu His
 275 280 285
 Val His Thr Ala Leu Phe Ser Ser Lys Asp Phe Ala Gly Lys Ser Gln
 290 295 300
 His Gly Val Tyr Gly Asp Ala Val Glu Glu Met Asp Trp Ser Val Gly
 305 310 315 320
 Gln Ile Leu Asn Leu Leu Asp Glu Leu Arg Leu Ala Asn Asp Ile Leu
 325 330 335
 Ile Tyr Phe Thr Ser Asp Gln Gly Ala His Val Glu Glu Val Ser Ser
 340 345 350
 Lys Gly Glu Ile His Gly Gly Ser Asn Gly Ile Tyr Lys Gly Gly Lys
 355 360 365
 Ala Asn Asn Trp Glu Gly Gly Ile Arg Val Pro Gly Ile Leu Arg Trp
 370 375 380

64

Pro Arg Val Ile Gln Ala Gly Gln Lys Ile Asp Glu Pro Thr Ser Asn
 385 390 395 400

Met Asp Ile Phe Pro Thr Val Ala Lys Leu Ala Gly Ala Pro Leu Pro
 405 410 415

Glu Asp Arg Ile Ile Asp Gly Arg Asp Leu Met Pro Leu Leu Glu Gly
 420 425 430

Lys Ser Gln Arg Ser Asp His Glu Phe Leu Phe His Tyr Cys Asn Ala
 435 440 445

Tyr Leu Asn Ala Val Arg Trp His Pro Gln Asn Ser Thr Ser Ile Trp
 450 455 460

Lys Ala Phe Phe Phe Thr Pro Asn Phe Asn Pro Val Gly Ser Asn Gly
 465 470 475 480

Cys Phe Ala Thr His Val Cys Phe Cys Phe Gly Ser Tyr Val Thr His
 485 490 495

His Asp Pro Pro Leu Leu Phe Asp Ile Ser Lys Asp Pro Arg Glu Arg
 500 505 510

Asn Pro Leu Thr Pro Ala Ser Glu Pro Arg Phe Tyr Glu Ile Leu Lys
 515 520 525

Val Met Gln Glu Ala Ala Asp Arg His Thr Gln Thr Leu Pro Glu Val
 530 535 540

Pro Asp Gln Phe Ser Trp Asn Asn Phe Leu Trp Lys Pro Trp Leu Gln
 545 550 555 560

Leu Cys Cys Pro Ser Thr Gly Leu Ser Cys Gln Cys Asp Arg Glu Lys
 565 570 575

Gln Asp Lys Arg Leu Ser Arg
 580

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 551 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Met Lys Ser Ala Pro Phe Leu Phe Leu Leu Gly Leu Leu Gly Leu Val
 1 5 10 15

65

Thr Ala Gln Thr Gln Asp Pro Ala Leu Leu Asp Leu Leu Arg Glu Asn
 20 25 30

Pro Asp Leu Leu Ser Leu Leu Leu Gln Ser Asn Glu His Arg Ala Pro
 35 40 45

Leu Val Lys Pro Asn Val Val Leu Leu Val Ala Asp Asp Met Gly Ser
 50 55 60

Gly Asp Leu Thr Ser Tyr Gly His Pro Thr Gln Glu Ala Gly Phe Ile
 65 70 75 80

Asp Lys Met Ala Ala Glu Gly Leu Arg Phe Thr Asn Gly Tyr Val Gly
 85 90 95

Asp Ala Val Cys Thr Pro Ser Arg Ser Ala Ile Met Ile Gly Arg Leu
 100 105 110

Pro Val Arg Ile Gly Thr Phe Gly Glu Thr Arg Val Phe Leu Pro Trp
 115 120 125

Thr Lys Thr Gly Leu Pro Lys Ser Glu Leu Thr Ile Ala Glu Ala Met
 130 135 140

Lys Glu Ala Gly Tyr Ala Ile Gly Met Val Gly Lys Trp His Leu Gly
 145 150 155 160

Met Asn Glu Asn Ser Ser Ile Asp Gly Ala His Leu Pro Phe Asn His
 165 170 175

Gly Phe Asp Phe Val Gly His Asn Leu Pro Phe Thr Asn Ser Trp Ser
 180 185 190

Cys Asp Asp Thr Gly Leu His Lys Asp Phe Pro Asp Ser Gln Arg Cys
 195 200 205

Tyr Leu Tyr Val Asn Ala Thr Leu Val Ser Gln Pro Tyr Gln His Lys
 210 215 220

Gly Leu Thr Gln Leu Phe Thr Asp Asp Ala Leu Gly Phe Ile Glu Asp
 225 230 235 240

Asn His Ala Asp Pro Phe Leu Tyr Val Ala Phe Ala His Met His
 245 250 255

Thr Ser Leu Phe Ser Ser Asp Asp Phe Ser Cys Thr Ser Arg Arg Gly
 260 265 270

Arg Tyr Gly Asp Asn Leu Leu Glu Met His Asp Ala Val Asp Lys Ile
 275 280 285

Val Asp Lys Leu Glu Glu Asn Asn Ile Ser Glu Asn Ile Ile Ile Phe
 290 295 300

Colo

Phe Ile Ser Asp His Gly Pro His Arg Glu Tyr Cys Glu Glu Gly Gly
 305 310 315 320

Asp Ala Ser Ile Phe Arg Gly Gly Lys Ser His Ser Trp Glu Gly Gly
 325 330 335

His Arg Ile Pro Tyr Ile Val Tyr Trp Pro Gly Thr Ile Ser Pro Gly
 340 345 350

Ile Ser Asn Glu Ile Val Thr Ser Met Asp Ile Ile Ala Ile Ala Ala
 355 360 365

Asp Leu Gly Gly Thr Thr Leu Pro Thr Asp Arg Ile Tyr Asp Gly Lys
 370 375 380

Ser Ile Lys Asp Val Leu Leu Glu Gly Ser Ala Ser Pro His Ser Ser
 385 390 395 400

Phe Phe Tyr Tyr Cys Lys Asp Asn Leu Met Ala Val Arg Val Gly Lys
 405 410 415

Tyr Lys Ala His Phe Arg Thr Gln Arg Val Arg Ser Gln Asp Glu Tyr
 420 425 430

Gly Leu Glu Cys Ala Gly Gly Phe Pro Leu Glu Asp Tyr Phe Asp Cys
 435 440 445

Asn Asp Cys Glu Gly Asp Cys Val Thr Glu His Asp Pro Pro Leu Leu
 450 455 460

Phe Asp Leu Met Arg Asp Pro Gly Glu Ala Tyr Pro Leu Glu Ala Cys
 465 470 475 480

Gly His Glu Asp Val Phe Leu Thr Val Lys Ser Thr Val Glu Glu His
 485 490 495

Lys Ala Ala Leu Val Lys Cys Thr Pro Leu Leu Asp Ser Phe Asp His
 500 505 510

Ser Ile Val Pro Cys Cys Asn Pro Ala Asn Cys Cys Ile Cys Asn Tyr
 515 520 525

Val His Glu Pro Gly Met Pro Glu Cys Tyr Gln Asp Gln Val Ala Thr
 530 535 540

Ala Ala Arg His Tyr Arg Pro
 545 550